



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00		<b>A2</b>	<b>(11) International Publication Number:</b> WO 00/58473 <b>(43) International Publication Date:</b> 5 October 2000 (05.10.00)
<b>(21) International Application Number:</b> PCT/US00/08621 <b>(22) International Filing Date:</b> 31 March 2000 (31.03.00)  <b>(30) Priority Data:</b> 60/127,607 31 March 1999 (31.03.99) US 60/127,636 2 April 1999 (02.04.99) US 60/127,728 5 April 1999 (05.04.99) US 09/540,763 30 March 2000 (30.03.00) US  <b>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</b> US 60/127,607 (CIP) Filed on 31 March 1999 (31.03.99) US 60/127,636 (CIP) Filed on 2 April 1999 (02.04.99) US 60/127,728 (CIP) Filed on 5 April 1999 (05.04.99) US 09/540,763 (CIP) Filed on 30 March 2000 (30.03.00)  <b>(71) Applicant (for all designated States except US):</b> CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).		<b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US).  <b>(74) Agent:</b> ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).  <b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>	
<b>(54) Title:</b> NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES, "ORFX"			
<b>(57) Abstract</b> <p>The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.</p>			

6322 SEP 24 2000

2775	95307987 (5549, 5550)	Novel Protein sim. GBank gi4689132 gb AAD2775.1 AF07704 - (AF077042), 30S ribosomal protein S7 homolog [Homo sapiens]	Contains protein domain (PF00177) - Ribosomal protein S7pSse	ribosomalprot	264488, 22278995, 56994075, 22278996, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331827, 29331828, 35696052, 265007, 21906754, 265017, 265019, 264448, 264682, 264369, 264288, 18108354, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 60170615, 264691, 18108370, 35696423, 65274791, 35695855, 264634, 60431850, 60170394, 56182323, 264558, 18108388, 22279000, 264563, 264565, 264486, 264567
2776	87791557 (5551, 5552)				56182575, 22278998, 22278999, 264259, 29331822, 29331824, 264908, 29331830, 264510, 33657402, 21906754, 5581386, 265017, 265019, 264448, 264288, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 265022, 60170615, 55810764, 55811576, 264555, 56526486, 22279000
2777	79818729 (5553, 5554)			UNCLASSIFIED	264907, 264766
2778	82112411 (5555, 5556)			UNCLASSIFIED	264907, 264593, 264760, 264628
2779	87649729 (5557, 5558)	Novel Protein sim. GBank gi4680711 gb AAD27745.1 AF13297 - (AF132970) CGI-36 protein [Homo sapiens]		UNCLASSIFIED	22278997, 264259, 29331824, 66714117, 35696052, 265006, 264512, 264448, 264288, 29148627, 18108364, 20281149, 18108370, 264629
2780	94679397 (5559, 5560)	Novel Protein sim. GBank gi4758524 ref NP_004825.1 PHGK1 - HPK/GCK-like kinase	Contains protein domain (PF00780) - CNH domain	kinase	29146499, 65274791, 264634, 264639
2781	91220057 (5561, 5562)	Novel Protein sim. GBank gi4469352 gb AAD212221 - (AF069502) ubiquitin specific protease UBP43 [Mus musculus]	Contains protein domain (PF00443) - Ubiquitin carboxy-terminal hydrolase family 2	ubiquitin	60424179, 29331824, 60424269, 66714117, 29331826, 56182435, 87166474, 265017, 264764, 56181562, 21906765, 21906766, 21906768, 35695917, 265020, 33657023, 35695855, 56182323, 87168518
2782	94233146 (5563, 5564)	Novel Protein sim. GBank gi4505013 ref NP_002310.1 PLRN1 - leucine-rich neuronal protein	Contains protein domain (PF00560) - Leucine Rich Repeat	struct	65274572, 22278996, 22278998, 60432049, 264259, 29331822, 29331824, 29331826, 60432289, 29331828, 264905, 264907, 264908, 264909, 52644045, 265009, 265017, 265018, 264804, 265019, 264760, 264683, 264288, 264766, 264685, 264686, 264768, 52644229, 264689, 21906768, 265020, 265021, 264691, 18108362, 264692, 33657023, 264693, 33657109, 33657349, 18108370, 264628, 263978, 35695855, 264557, 56182323, 83373044, 18108385

gttcatgaca tcttctatcc cttcccccaa tcagagggag aggacgagct ctgctttctc  
 1860  
 cgcgccaatg aatgcaagac cggtttctgc catttgtaga aagtcaccgc cgttttaaaa  
 1920  
 tcccagggct acgattggag tgagcccttc agccccgggg aagatgaatt taagtgtccc  
 1980  
 attaaggaag agattgctct gaccagcggg gaatgggagg ttttggcgag gcacggctcc  
 2040  
 aagatctggg tcaatgagga gaccaagctg gtgtacttcc agggcaccaa ggacacgccg  
 2100  
 ctggagcacc acctctacgt ggtagctat gaggcggccg gcgagatcgt acgcctcacc  
 2160  
 acgcccggct tctcccatag ctgctccatg agccagaact tcgacatgtt cgtcagccac  
 2220  
 tacagcagcg tgagcacgcc gccctgcgtg cacgtctaca agctgagcgg ccccgacgac  
 2280  
 gacccccctgc acaagcagcc ccgcttctgg gctagcatga tggaggcagc cagctgtccc  
 2340  
 ccggattatg ttctccaga gatcttccat ttccacacgc gctcggatgt gcggctctac  
 2400  
 ggcattgatc acaagcccca cgccttgtag cacatcacca aaaaatctac cgtcttcgag  
 2460  
 a  
 2461

&lt;210&gt; 2780

&lt;211&gt; 720

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 2780

Met	His	Ser	Glu	Gln	Glu	Gly	Gln	His	Val	Gln	Arg	Pro	Cys	Gly	Gly
1				5					10					15	
Lys	Glu	Phe	Gly	Leu	Phe	Glu	Glu	Leu	Ser	Glu	Gly	Ser	Phe	Gly	Trp
			20					25					30		
Val	Thr	Gly	Ile	Arg	Arg	Met	Arg	Phe	Lys	Gly	Leu	Ala	Gly	Val	Asp
		35					40					45			
Ser	Ser	Leu	Glu	Val	Val	Ser	Leu	Leu	Pro	Pro	Arg	Ser	Phe	Ser	Leu
		50				55					60				
Asn	Ser	Glu	Gly	Ala	Glu	Arg	Met	Ala	Thr	Thr	Gly	Thr	Pro	Thr	Ala
65					70					75				80	
Asp	Arg	Gly	Asp	Ala	Ala	Ala	Thr	Asp	Asp	Pro	Ala	Ala	Arg	Phe	Gln
			85					90					95		
Val	Gln	Lys	His	Ser	Trp	Asp	Gly	Leu	Arg	Ser	Ile	Ile	His	Gly	Ser
		100						105					110		
Arg	Lys	Tyr	Ser	Gly	Leu	Ile	Val	Asn	Lys	Ala	Pro	His	Asp	Phe	Gln
		115					120					125			
Phe	Val	Gln	Lys	Thr	Asp	Glu	Ser	Gly	Pro	His	Ser	His	Arg	Leu	Tyr
		130				135					140				
Tyr	Leu	Gly	Met	Pro	Tyr	Gly	Ser	Arg	Glu	Asn	Ser	Leu	Leu	Tyr	Ser
145					150					155				160	
Glu	Ile	Pro	Lys	Lys	Val	Arg	Lys	Glu	Ala	Leu	Leu	Leu	Leu	Ser	Trp
				165					170					175	
Lys	Gln	Met	Leu	Asp	His	Phe	Gln	Ala	Thr	Pro	His	His	Gly	Val	Tyr

180 185 190  
 Ser Arg Glu Glu Leu Leu Arg Glu Arg Lys Arg Leu Gly Val Phe  
 195 200 205  
 Gly Ile Thr Ser Tyr Asp Phe His Ser Glu Ser Gly Leu Phe Leu Phe  
 210 215 220  
 Gln Ala Ser Asn Ser Leu Phe His Cys Arg Asp Gly Gly Lys Asn Gly  
 225 230 235 240  
 Phe Met Val Ser Pro Gly Pro Gly Cys Val Ser Pro Met Lys Pro Leu  
 245 250 255  
 Glu Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys Ile Cys  
 260 265 270  
 Pro Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp Leu Trp  
 275 280 285  
 Val Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe Cys His  
 290 295 300  
 Gln Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly Val Ala  
 305 310 315 320  
 Thr Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp  
 325 330 335  
 Cys Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg  
 340 345 350  
 Ile Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val  
 355 360 365  
 Pro Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro  
 370 375 380  
 Arg Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe  
 385 390 395 400  
 Gln Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu  
 405 410 415  
 Val Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg  
 420 425 430  
 Ala Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp  
 435 440 445  
 Arg Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe  
 450 455 460  
 Ile Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala  
 465 470 475 480  
 Val Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn  
 485 490 495  
 Val Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu  
 500 505 510  
 Gly Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly  
 515 520 525  
 Phe Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr  
 530 535 540  
 Asp Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys Cys Pro  
 545 550 555 560  
 Ile Lys Glu Glu Ile Ala Leu Thr Ser Gly Glu Trp Glu Val Leu Ala  
 565 570 575  
 Arg His Gly Ser Lys Ile Trp Val Asn Glu Glu Thr Lys Leu Val Tyr  
 580 585 590  
 Phe Gln Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr Val Val  
 595 600 605  
 Ser Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro Gly Phe

610	615	620
Ser His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val Ser His		
625	630	635
Tyr Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys Leu Ser		
645	650	655
Gly Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp Ala Ser		
660	665	670
Met Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro Glu Ile		
675	680	685
Phe His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met Ile Tyr		
690	695	700
Lys Pro His Ala Leu Gln His Ile Thr Lys Lys Ser Thr Val Phe Glu		
705	710	715
		720

<210> 2781  
 <211> 1268  
 <212> DNA  
 <213> Homo sapiens

<400> 2781  
 gtcgacggac ttcaggaagt gcagcgccag gcacaagagg ggaagaatat aggcaccacc  
 60  
 aagaagggaa tcggaccaac ctactcttcc aaagctgccc ggacaggcct ccgcatctgc  
 120  
 gacctctgt cagatcttga tgagtttct tccagattca agaacctggc ccaccagcac  
 180  
 cagtcgatgt tccccaccct ggaaatagac attgaaggcc aactcaaaag gctcaagggc  
 240  
 tttgctgagc ggatcagacc catggtccga gatggtgttt actttatgta tgaggcactc  
 300  
 cacggccccc ccaagaagat cctggtggag ggtgccaacg ccgccctcct cgacattgac  
 360  
 ttcgggacct acccctttgt gacttcatcc aactgcaccg tgggcggtgt gtgcacgggc  
 420  
 ctgggcatcc cccgcagaa cataggtgac gtgtatggcg tggtgaaagc ctataccaca  
 480  
 cgtgtgggca tcggggcctt cccaccgag cagatcaacg agattggagg cctgctgcag  
 540  
 acccgcgcc acgagtgggg agtgaccaca ggcaggaaga ggcgctgcgg ctggctcgac  
 600  
 ctgatgattc taagatatgc tcacatggtc aacggattca ctgcgctggc cctgacgaag  
 660  
 ctggacatcc tggacgtact ggggtgaggtt aaagtcggtg tctcatacaa gctgaacggg  
 720  
 aaaaggattc cctatttccc agctaaccag gagatgcttc agaaggctga agttgagtat  
 780  
 gaaacgctgc ctgggtggaa agcagacacc acaggcgcca ggaggtggga ggacctgccc  
 840  
 ccacaggccc agaactacat ccgctttgtg gagaatcacg tgggagtcgc agtcaaattg  
 900  
 gttggtgttg gcaagtcaag agagtcgatg atccagctgt tttagtcgca gactgagctg  
 960  
 atcccaacag gccctggcag cgtctggact tgtgtaaaca gcagcagtca cgttcctcgg  
 1020